Supplementary Material

Expert assessment on volumetric laser endomicroscopy full-scans in Barrett’s esophagus patients with or without high-grade dysplasia or early cancer

Maarten Struyvenberg, Allon Kahn, David Fleischer et al.

Appendix 1s Methods

Scoring of established VLE features

Reviewers assessed full-scan videos using a methodology at their discretion, based on their expertise in scoring and evaluating VLE scans. They were asked to rate the dysplastic videos using established dysplastic VLE features. A combination of the following previously published scoring systems was used: VLE-DA [1], OCT-SI [2], and the VLE-prediction score [3]. According to these studies: absence of mucosal layering, higher surface intensity compared to subsurface intensity, and presence of irregular glands, produced the highest likelihood for neoplasia.

Effect of “poor performing assessors” or “difficult videos”?

For the influence of poor performing assessors, we deleted the results of the four assessors with the worst performance for overall accuracy, correct neoplastic diagnosis, and lesion localization. To analyze the influence of “difficult videos”, we deleted the three worst-scored neoplastic videos.
Appendix 2s Methods: Statistical analysis

Calculation of diagnostic accuracy, sensitivity and specificity

We calculated accuracy, sensitivity, and specificity for each assessor separately according to the following formula.

<table>
<thead>
<tr>
<th>Assessors diagnoses</th>
<th>Ground Truth</th>
</tr>
</thead>
<tbody>
<tr>
<td>neoplastic</td>
<td>neoplastic</td>
</tr>
<tr>
<td></td>
<td>True positive (TP)</td>
</tr>
<tr>
<td></td>
<td>False positive (FP)</td>
</tr>
<tr>
<td>non-dysplastic</td>
<td>False negative (FN)</td>
</tr>
<tr>
<td></td>
<td>True negative (TN)</td>
</tr>
</tbody>
</table>

**Sensitivity** = Proportion of correctly identified neoplastic patients = TP/(TP+FN)

**Specificity** = Proportion of correctly identified non-dysplastic patients = TN/(TN+FP)

**Accuracy** = Proportion of correct diagnoses from the total number of full scan videos = (TN+TP)/ (TN+TP+FP+FN)
Appendix 3 Results: Secondary outcome measurements

Association between established VLE features and neoplasia

Neoplastic diagnosis was primarily guided by the VLE features: lack of layering (87%), higher surface intensity compared to subsurface intensity (74%), and presence of > 5 irregular glands or ducts (77%). When the expert localized the most abnormal VLE area outside the true lesion location, lack of layering, higher surface intensity compared to subsurface intensity and presence of > 5 irregular glands were deemed to be present in 76%, 59% and 49%, respectively. For the non-dysplastic videos incorrectly labelled as neoplastic, the incorrect diagnosis was primarily guided by the VLE features ‘lack of layering’ in 76% and ‘higher surface compared to subsurface intensity’ in 75%.

Appendix 4 Results: Secondary outcome measurements

Effect of “poor performing assessors” or “difficult videos”? 

For the influence of poor performing assessors, we deleted the results of the four assessors with the worst performance for overall accuracy, correct neoplastic diagnosis, and lesion localization. The remaining VLE-experts correctly labelled 77% (before 73%) of neoplastic VLE-videos as “neoplastic”. NDBE videos were correctly identified in 54% (before 52%), and lesion localization was 57% (before 54%). To analyze the influence of “difficult videos”, we deleted the three worst-scored neoplastic videos. For the remaining neoplastic videos, a correct diagnosis was reached in 80% (before 73%), and lesion localization was 63% (before 54%). Additionally, we analyzed the four best-scored neoplastic videos. No differences were found for type of neoplasia, tumor depth of invasion, or BE length, and correct neoplastic diagnosis in the worst and best scored videos. Fig. 1s shows these neoplastic videos with corresponding clinical information, assessor performance and specific VLE features.
References supplementary materials


**Supplementary Figure 1.** A) Four endoscopic overview images of neoplastic BE with corresponding VLE images with the best assessor performance. Endoscopically visible neoplasia is marked with a blue line in the adjacent endoscopic image. The corresponding area on the VLE image is marked using a red line.
Supplementary Figure 1. B) Three endoscopic overview images of neoplastic BE with corresponding VLE images with the least accurate assessor performance. Endoscopically visible neoplasia is marked with a blue line in the adjacent endoscopic image. The corresponding area on the VLE image is marked using a red line.